


# Analiza danych eksperymentalnych wspomagana komputerowo

KIEROWNIK  
STUDIÓW DOKTORANCKICH  
Wydział Elektroniki WAT  
dr hab. inż. Andrzej DOBROWOLSKI

## Informacje ogólne

2014 -06- 2 7

Kod przedmiotu:	WELXXCD-ADE	Kod Erasmus:	(brak danych)
Nazwa przedmiotu:	Analiza danych eksperymentalnych wspomagana komputerowo		
Jednostka:	Wydział Elektroniki		
Grupy:			
Punkty ECTS i inne:	3.00 		
Język prowadzenia:	polski		
Rodzaj studiów:	III stopnia		
Rodzaj przedmiotu:	obowiązkowy		
Forma zajęć liczba godzin/rygor:	W 16/x, Ć 14		
Przedmioty wprowadzające:	Matematyka. Wymagania wstępne: znajomość rachunku macierzowego i rachunku prawdopodobieństwa.		
Programy:	Dyscyplina naukowa studiów: Elektronika, Telekomunikacja		
Autor:	dr hab. inż. Jacek Jakubowski		
Skrócony opis:	<p>Prezentowane w ramach przedmiotu treści obejmują wnioskowanie statystyczne jako podstawową metodologię tradycyjnej analizy danych oraz wybrane metody analizy eksploracyjnej, której celem jest wykrycie nieznanych prawidłowości, związków i anomalii w danych. W ramach zajęć audytoryjnych omawiane są zagadnienia estymacji parametrów modeli danych, modeli regresji dla przypadków homo i heteroskedastycznego, zagadnienia graficznej reprezentacji danych jedno i wielowymiarowych, współzmienności danych, przetwarzania wyników eksperymentu za pomocą technik wielowymiarowej analizy danych PCA i LDA oraz zagadnienia dyskryminacji, obejmujące problemy liniowo i nieliniowo separowalne z wykorzystaniem prostych metod wytwarzania hiperpłaszczyzn rozdzielających, metod klasyfikacji minimalnoodległościowej, najbliższych sąsiadów, k średnich i grupowania hierarchicznego. Omawiane na zajęciach audytoryjnych treści są utrwalane na ćwiczeniach rachunkowych realizowanych w oparciu o zbiór procedur Matlaba.</p>		
Pełny opis:	<p>Wykłady realizowane są w formie werbalno-wizualnej prezentacji następujących treści:</p> <ol style="list-style-type: none"><li>1. Zagadnienia wprowadzające Zasady realizacji i zaliczania przedmiotu. Cele i podział metod analizy danych. Charakterystyka narzędzi programistycznych do analizy danych: Statistica, Matlab Statistics Toolbox, pakiet R. Opis danych jednowymiarowych. Zastosowanie wybranych metod wnioskowania statystycznego w analizie danych.</li><li>3. Wprowadzenie do analizy regresji Założenia estymacji parametrów modelu regresji. Rachunek skalarny modelu regresji liniowej. Rachunek macierzowy. Przypadek heteroskedastyczny. Macierz kowariancji estymatora parametrów. Regresja wielomianowa i krzywoliniowa.</li><li>4. Opis danych wielowymiarowych Kowariancja jako miara współzmienności. Współczynnik korelacji liniowej Pearsona. Ilościowe znaczenie współczynnika korelacji. Reprezentacja graficzna danych wielowymiarowych: wykresy rozproszeń, gwiazdowe,</li></ol>		

graficzna danych wielowymiarowych: wykresy rozproszeń, gwiazdowe, twarzy Chernoffa, Andrews. Opis matematyczny operacji rzutowania punktu na wyróżniony kierunek.

#### 5. Transformacja PCA

Macierz kowariancji. Idea przekształcenia PCA. Dekompozycja macierzy kowariancji na wektory i wartości własne. Własności macierzy przekształcenia i danych w przestrzeni docelowej.

#### 5. Transformacja LDA

Kryterium transformacji LDA. Przebieg transformacji dla wariantu dwuklasowego. Schemat transformacji dla wariantu wieloklasowego.

#### 6. Wprowadzenie do analizy dyskryminacji

Podział metod. Klasyfikacja wzorcowa w problemie dwu klas separowalnych liniowo. Metody wyznaczania hiperpłaszczyzn rozdzielających. Przypadek wieloklasowy.

#### 7. Problemy nieseparowalne liniowo

Miary odległości w analizie danych. Klasyfikacja minimalno-odległościowa. Metoda najbliższych sąsiadów (k-NN).

#### 8. Klasyfikacja bezwzorcowa

Metoda k-średnich i metoda grupowania hierarchicznego.

Omawiane na zajęciach audytoryjnych treści są utrwalane na ćwiczeniach rachunkowych, które obejmują następujące treści:

#### 1. Rozkłady wyników eksperymentu

Wykorzystanie środowiska Matlab do tworzenia raportu z analizy danych na przykładzie empirycznego wyznaczania funkcji gęstości prawdopodobieństwa. Obliczanie miar położenia, rozrzutu i kształtu rozkładów danych. Miary opisu wyników zawierających dane odstające. Reprezentacja graficzna danych.

#### 2. Oszacowanie współczynników modeli regresji

Konstrukcja macierzy regresorów. Obliczenia współczynników prostej aproksymującej dyskretne wyniki eksperymentów. Wyznaczanie wariancji parametrów modelu. Badanie stopnia dopasowania wielomianu do danych eksperymentalnych.

#### 3. Badanie korelacji i wizualizacja danych wielowymiarowych

Obliczenia współczynnika korelacji Pearsona. Badanie istotności korelacji. Metody zobrazowania danych wielowymiarowych.

#### 4. Badanie własności transformacji PCA

Obliczenia macierzy kowariancji. Wyznaczanie macierzy przekształcenia PCA. Porównanie danych z przestrzeni oryginalnej i przestrzeni docelowej. Redukcja wymiaru danych. Zastosowanie PCA do stratnej kompresji obrazów.

#### 5. Przykłady transformacji danych wielowymiarowych za pomocą LDA

Przykłady rozwiązywania zadań transformacji LDA dla danych dwuwymiarowych w wariacie dwuklasowym. Zastosowanie do analizy przypadku wielowymiarowego i wieloklasowego.

#### 6. Hiperpłaszczyzny rozdzielające

Przykłady obliczeń w wariacie dwuklasowym z wykorzystaniem metody reprezentantów klas i metody aproksymacji. Implementacja modelu neuronu Rosenblatta.

#### 7. Analiza skupień

Przykłady zastosowań metod klasteryzacji. Tworzenie prezentacji wideo wyników analizy danych.

Literatura: Literatura podstawowa

1. J. Koronacki, J. Ćwik, Statystyczne systemy uczące się, Akademicka Oficyna Wydawnicza EXIT, wyd. 2, 2008.
2. W. Kwiatkowski, Metody automatycznego rozpoznawania wzorców, Instytut Automatyki i Robotyki Wydziału Cybernetyki WAT, wyd. 1, 2001.
3. M. Dobosz, Wspomagana komputerowo statystyczna analiza wyników badań, Akademicka Oficyna Wydawnicza EXIT, 2001.

#### Literatura uzupełniająca

uzupełniająca:

1. W. Klonecki, Statystyka dla inżynierów, Wydawnictwo Naukowe PWN, wyd. 1, 1999.
2. J. R. Taylor, Wstęp do analizy błędów pomiarowych, Wydawnictwo Naukowe PWN, wyd. 1 - 1995, wyd. 2 - 1999.

**Efekty uczenia:** ADE\_W1 / Doktorant ma poszerzoną i pogłębioną wiedzę w zakresie zastosowania narzędzi wnioskowania statystycznego jako podstawowej metodologii tradycyjnej analizy danych eksperymentalnych / EiT\_W01

ADE\_W2 Doktorant zna i rozumie algorytmy eksploracyjnej analizy danych wielowymiarowych (data mining) ukierunkowane na wizualizację, redukcję wymiarowości, ekstrakcję cech charakterystycznych, predykcję, klasyfikację i analizę skupień / EiT\_W07

ADE\_U1 Doktorant potrafi przygotować i przeprowadzić eksperymenty numeryczne z wykorzystaniem poznanych metod wielowymiarowej analizy danych eksperymentalnych / EiT\_U06

ADE\_U2 Doktorant potrafi opracować szczegółową dokumentację przeprowadzonej analizy danych z wykorzystaniem narzędzi wytwarzania wersji elektronicznej raportu i narzędzi przygotowania elementów prezentacji multimedialnej oraz zawierającą omówienie uzyskanych wyników / EiT\_U03

ADE\_K1 Doktorant potrafi kreatywnie myśleć przy rozwiązywaniu problemu badawczego oraz współdziałać i pracować w małym zespole / EiT\_K03, EiT\_K04

**Kryteria oceniania:** Warunkiem zaliczenia przedmiotu jest uzyskanie pozytywnych ocen ze wszystkich raportów z analizy danych wykonywanych na ćwiczeniach oraz uzyskanie pozytywnej oceny z egzaminu, który odbędzie się w formie pisemnej i ustnej.

Efekty ADE\_W1, ADE\_W2 sprawdzane są na: ćwiczeniach rachunkowych i egzaminie.

Efekty ADE\_U1, ADE\_U2 i ADE\_K1 sprawdzane są na: ćwiczeniach rachunkowych oraz na podstawie sporządzanych przez studentów raportów z analizy danych.

autor sylabusa

*Jacek Jakubowski*

dr hab. inż. Jacek Jakubowski

DYREKTOR  
Instytutu Systemów Elektronicznych  
Wydziału Elektroniki WAT

*J. Dąbrowski*  
dr hab. inż. Jacek DĄBROWSKI